

3/5/1

3/5/1

ALOG(R) File 351:Derwent WPI

2004 Thomson Derwent. All rts. reserv.

2628123 \*\*Image available\*\*

: Acc No: 1999-434227/199937

PX Acc No: N99-323556

voice recognition device for telephone circuit - uses voice recognition  
 host of downloading unit to provide voice recognition software module to  
 telephone terminal so that telephone terminal can respond to voice  
 recognition result via circuit interface

: Invent Assignee: NEC CORP (NIDE ); KATO G (KATO-I)

: Inventor: KATO G

: Number of Countries: 002 Number of Patents: 004

: Invent Family:

Invent No	Kind	Date	Applicat No	Kind	Date	Week
11175091	A	19990702	JP 97352210	A	19971205	199937 B
3055514	B2	20000626	JP 97352210	A	19971205	200035
20020094066	A1	20020718	US 98205242	A	19981204	200254
6522725	B2	20030218	US 98205242	A	19981204	200317

: Priority Applications (No Type Date): JP 97352210 A 19971205

: Invent Details:

Invent No	Kind	Lan	Pg	Main IPC	Filing Notes
11175091	A		6	G10L-003/00	
3055514	B2		6	G10L-015/00	Previous Publ. patent JP 11175091
20020094066	A1			H04M-001/64	
6522725	B2			H04M-001/64	

: Abstract (Basic): JP 11175091 A

NOVELTY - A telephone terminal (3) responds to the voice recognition result via a circuit interface (31) using a predetermined format, in which a voice-recognition software module is downloaded from the voice recognition host (11) of a station installation type downloading unit (1) via a telephone network (2). DETAILED DESCRIPTION - The downloading apparatus has a controller (14) that regulates the input data from a network control unit (NCU) (15) which controls the predetermined call from the telephone network. An audio response circuit (12) provides a voice responds depending on the demand from the voice recognition host. A database (13) sequentially provides predetermined data based on the demand from the voice recognition host. The telephone terminal has a calculation circuit (33) which executes the voice recognition software received and stored by a memory (32), to perform a voice recognition.

USE - For telephone circuit.

ADVANTAGE - Enables flexible service modification in a station when building a voice recognition system, by replacing voice-recognition software module. DESCRIPTION OF DRAWING(S) - The figure shows the component diagram of a voice recognition device. (1) Downloading unit; (2) Telephone network; (3) Telephone terminal; (11) Voice recognition host; (12) Audio response circuit; (13) Database; (14) Controller; (15) NCU; (31) Circuit interface; (32) Memory; (33) Calculation circuit.

Dwg.1/4

le Terms: VOICE; RECOGNISE; DEVICE; TELEPHONE; CIRCUIT; VOICE; RECOGNISE  
 HOST; UNIT; VOICE; RECOGNISE; SOFTWARE; MODULE; TELEPHONE; TERMINAL; SO  
 TELEPHONE; TERMINAL; CAN; RESPOND; VOICE; RECOGNISE; RESULT; CIRCUIT;  
 NTERFACE

: Invent Class: P86; T01; W01; W04

: International Patent Class (Main): G10L-003/00; G10L-015/00; H04M-001/64

: International Patent Class (Additional): G06F-009/445; G10L-015/28;

: 04B-003/46; H04B-017/00; H04M-003/42

**THIS PAGE BLANK (USPTO)**



**THIS PAGE BLANK (USPTO)**

JPA11-17509/ which corresponds to USP 6,522,725

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平11-175091

(43) 公開日 平成11年(1999) 7月2日

(51) Int. Cl. <sup>6</sup>  
G10L 3/00  
G06F 9/445  
H04M 3/42

識別記号  
551

F I  
G10L 3/00 551 A  
H04M 3/42 P  
G06F 9/06 420 J

審査請求 有 請求項の数 8 F D (全 6 頁)

(21) 出願番号 特願平9-352210

(22) 出願日 平成9年(1997)12月5日

(71) 出願人 000004237

日本電気株式会社  
東京都港区芝五丁目7番1号

(72) 発明者 加藤 剛

東京都港区芝五丁目7番1号 日本電気株式会社内

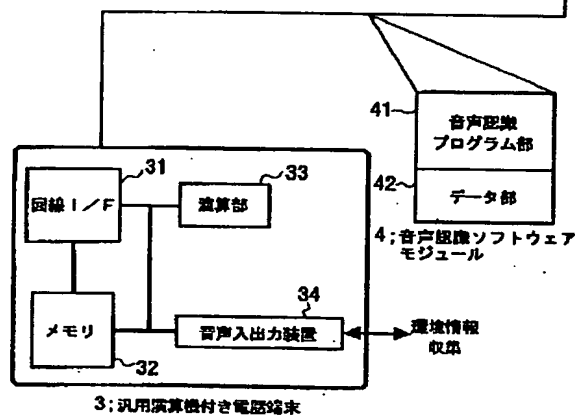
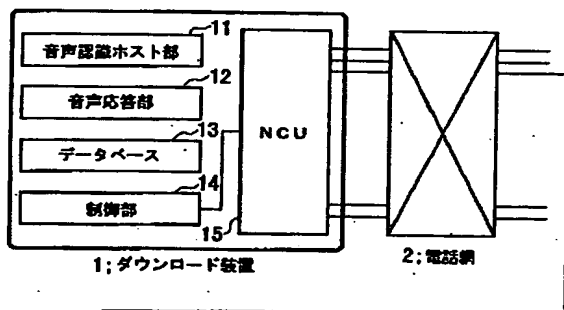
(74) 代理人 弁理士 加藤 朝道

(54) 【発明の名称】 電話回線用音声認識装置

(57) 【要約】

【課題】 端末型の音声認識性能を備え局側での柔軟なサービス変更が出来るようにした電話回線用音声認識装置の提供。

【解決手段】 回線制御装置からの入力データを制御する制御部、制御部を通して認識結果を受け取りサービスを実行する音声認識ホスト部、音声認識ホスト部からの要求に応じて音声応答する音声応答部、音声認識ホスト部からの要求に応じデータを提供するデータベースを備えた局設置ダウンロード装置と、ダウンロード装置から電話網を通してダウンロードされる音声認識ソフトウェアモジュールを受信しこれを格納する記憶部、該音声認識ソフトウェアモジュールを実行し音声認識を行う演算部、音声入出力部、を備え音声認識結果を音声認識ホスト部で規定された形式で回線インタフェースを介して応答する電話端末と、を備える



## 【特許請求の範囲】

【請求項1】電話網からの通話を制御する回線制御手段と、

前記回線制御手段からの入力データを制御する前記制御手段と、

前記制御手段を通して認識結果を受け取りサービスを実行する音声認識ホスト手段と、

前記音声認識ホスト手段からの要求に応じて音声で応答する音声応答手段と、

前記音声認識ホスト手段からの要求に応じ順次データを10 提供するデータベースと、を備えた局設置型のダウンロード装置と、

前記ダウンロード装置から前記電話網を通してダウンロードされる音声認識ソフトウェアモジュールを回線インタフェースで受信しこれを格納する記憶手段と、

前記記憶手段に格納された音声認識ソフトウェアモジュールを実行し音声認識を行う演算手段と、

音声入出力手段と、

を備え、音声認識結果を前記ダウンロード装置の前記音声認識ホスト手段で規定された形式で前記回線インタフ20 ェースを介して応答する電話端末と、

を備えたことを特徴とするような電話回線用音声認識装置。

【請求項2】前記電話端末側に、話者の話者性、すなわち発話タイミング、発話音量、発話の癖などを示すパラメータを保存するか、もしくは、前記音声認識ホスト手段で話者が特定できた時点で、前記音声認識ソフトウェアモジュールに前記パラメータが付加されて前記電話端末にダウンロードされ、さらに端末種別データが前記音声認識ソフトウェアモジュールに付加されてダウンロード30 される、ことを特徴とする請求項1記載の電話回線用音声認識装置。

【請求項3】前記音声認識ソフトウェアモジュールが、音声認識用プログラムよりなる音声認識プログラム部と、前記音声認識プログラム部に付随するデータ部とを含むことを特徴とする請求項1記載の電話回線用音声認識装置。

【請求項4】前記音声認識ソフトウェアモジュールに回線の状態をテストするためのデータを付加して前記電話端末にダウンロードし、前記電話端末はこのテスト用データを用いて回線特定の試験を行う、ことを特徴とする40 請求項1記載の電話回線用音声認識装置。

【請求項5】前記ダウンロード装置が、前記音声認識ソフトウェアモジュールに替わって、所望のサービスを前記電話端末側で実行するためのソフトウェアモジュールをダウンロードする、ことを特徴とする請求項1記載の電話回線用音声認識装置。

【請求項6】局側のダウンロード装置から電話網を通してダウンロードされる音声認識ソフトウェアモジュールを回線インタフェースで受信しこれを格納する記憶手段50

と、

前記記憶手段に格納された音声認識ソフトウェアモジュールを実行して音声認識を行う演算手段と、

音声入出力手段と、

を備えたことを特徴とするような電話端末。

【請求項7】局側から音声認識用ソフトウェアモジュールを電話回線を通して電話端末にダウンロードし、前記電話端末が、前記ダウンロードされた前記音声認識用ソフトウェアモジュールを処理装置で実行することで加入者の音声を確認し、該認識結果を前記局側で予め規定された形式で前記電話回線を通して転送する手段を備え、

前記局側が、前記電話端末側から転送された認識結果に基づき所定のサービスシナリオを実行する手段を備え、サービスを前記加入者に供給する、ことを特徴とする電話回線用音声認識装置。

【請求項8】前記局側に、前記サービスを実行する手段の要求に応じて音声で応答する手段を更に備えたことを特徴とする請求項7記載の電話回線用音声認識装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、電話回線用音声認識装置に関し、特に、音声品質の劣化により認識率の低下を回避し、且つ局側でのサービス変更に柔軟に対処可能な電話回線用音声認識装置に関する。

【0002】

【従来の技術】従来の電話回線用音声認識装置は、局（センター）設置型か、又は端末型かに分かれている。図3に、局側に音声認識装置を設置した従来の構成の一例をブロック図にて示す。図3を参照すると、電話端末301、電話網310、音声認識ユニット内蔵交換機320からなり、電話端末301は、回線インタフェース303及び音声入力装置302を備え、音声認識ユニット内蔵交換機320は、音声認識部321、音声応答部322、データベース323、制御部324、NCU（回線制御装置）325を備えている。この局設置型音声認識方式においては、回線状態により音声品質の劣化し（例えば漏話等の雑音、信号歪）、センター側の音声認識部321に入力される音声信号の品質が悪くなり、このため十分な認識性能を得るためには、処理内容、および音声データ認識用のデータベースのデータ量が増大することになる。また、センター側では不特定話者の音声を認識するために、高度な認識技術を実装することが必要とされている。

【0003】この種の問題の解決を図るものとして、例えば特開昭58-76890号公報には、端末側に音声認識用の非線形時間軸照合用部LSIを設け、端末での音声認識結果であるコード信号を電話回線を通してセンターに転送する音声認識装置が提案されている。図4に、この従来の方式をブロック図にて示す。図4を参照する

と、電話端末401は、音声入出力装置402、回線インタフェース403、音声認識装置404を備え、音声認識装置404は、例えばデジタル信号に変換された音声データを内蔵する標準パタンの照合し、最適に照合したものを認識結果とし、認識結果を回線インタフェース401を介してセンターに送る。

【0004】しかしながら、この従来方式では、音声認識機能は始めから端末側に固定されており、局側との連携したサービス変更等が困難であり、センター側のアプリケーション（サービス）の仕様変更の自由度が縮減されてしまうことになり、さらに端末側の音声認識装置の性能向上も容易ではなかった。

【0005】

【発明が解決しようとする課題】上記したように、従来の電話回線用音声認識装置は、局（センター）設置型か、端末型かのいずれかに分かれており、第1の問題点として、局設置型音声認識方式においては、回線状態による品質劣化をともなった音声の認識に対応した処理及び構成が必要とされ、装置の大型化、コスト上昇を招くという問題を有している。その理由は、局側の音声認識装置には、サービス利用者の音声回線を通して入力されるためである。

【0006】第2の問題点は、例えば上記特開昭58-76890号公報等に提案される端末型音声認識方式は、局（センター）と一体となったサービスの提供、及び最新のサービスの提供等は考慮されていない、ということである。その理由は、端末型音声認識方式において音声認識機能は、端末に組み込み込まれた機能に依存し、局側との動的なプログラム交換機能が実現されていないためである。

【0007】したがって、本発明は、上記問題点に鑑みてなされたものであって、その目的は、端末型の音声認識性能を備え、さらに局側での柔軟なサービス変更が出来るようにした電話回線用音声認識装置を提供することにある。

【0008】

【課題を解決するための手段】前記目的を達成するため、本発明の電話回線用音声認識装置は、局側から音声認識用ソフトウェアモジュールを電話回線を通して電話端末にダウンロードし、前記電話端末は、前記ダウンロードされた前音声認識用ソフトウェアモジュールを処理装置で実行することで加入者の音声認識し、認識結果を前記局側で予め規定された形式で前記電話回線を通して転送する手段を備え、前記局側に前記認識結果に基づき所定のサービスを実行する手段を備え、サービスを前記加入者に供給する、ことを特徴とする。

【0009】

【発明の実施の形態】本発明の実施の形態について以下に説明する。本発明の電話回線用音声認識装置は、その好ましい実施の形態において、音声認識部を電話端末側

と局側の2つに分け、演算手段で実行されることで音声認識機能を実現するプログラム（図1の音声認識ソフトウェアモジュール4）を電話端末側に置き、その認識結果を利用して動作する手段（図1の音声認識ホスト部11）を局側（図1のダウンロード装置1内）に配置するようにしたものである。

【0010】より詳細には、本発明の実施の形態においては、音声認識ソフトウェアモジュールを局側から電話端末にダウンロードする方式を取り、局側の認識ホストの仕様変更などに迅速に対応出来るようにしたものであり、電話網からの通話を制御して制御手段に接続する回線制御装置（図1の15）と、この回線制御装置からの入力データを制御する制御手段（図1の14）と、制御手段を通して認識結果を受け取りサービスを実行する音声認識ホスト手段（図1の11）と、音声認識ホスト手段からの要求に応じて音声で応答する音声応答手段（図1の12）と、音声認識ホスト手段からの要求に応じ順次データを提供するデータベース（図1の14）と、を備えた局設置型のダウンロード装置（図1の1）を備えている。

【0011】そして、本発明の実施の形態において、電話端末（図1の3）は、このダウンロード装置から電話網（図1の2）を通してダウンロードされる音声認識ソフトウェアモジュール（図1の4）を回線インタフェース手段（図1の32）で受信しこれを格納する記憶手段（図1の31）と、この記憶手段に格納された音声認識ソフトウェアモジュールを実行し音声認識を行う演算処理手段（図1の33）と、音声入出力手段（図1の34）と、を備え、音声認識結果を前記ダウンロード装置の音声認識ホスト手段で規定された形式で前記回線インタフェース手段を介して応答する。

【0012】このように、本発明の実施の形態においては、音声認識用のプログラムを局側と電話端末側の2つにわけ、電話端末側の音声認識プログラムは局側から供給する構成としたことにより、音声認識プログラムを局側で一元的に管理でき、保守性を向上する。特に、局設置型方式の長所である、利用者に対するサービスの入替えが容易になる。

【0013】また、局側から電話端末側に、実質的な音声認識用プログラムを置き、電話端末側で音声認識を行うことから、局設置型音声認識方式に不可避の問題であった音声の回線劣化による認識率低下を回避することができる。

【0014】このように、本発明の実施の形態によれば、局設置型、及び端末型の電話回線用音声認識装置の双方の長所を兼ね備えるものである。

【0015】

【実施例】上記した本発明の実施の形態について更に詳細に説明すべく、本発明の実施例について図面を参照して以下に詳細に説明する。

10

20

30

40

50

【0016】図1は、本発明の一実施例の構成を示す図である。図1において、1は電話交換機に併設された局側のダウンロード装置であり、音声認識ホスト部11、音声応答部12、データベース13、制御部14、回線制御部(NCU)15を備えている。2は電話網、3は電話端末であり、電話端末(「汎用演算器つき電話端末」ともいう)3は、回線インタフェース31、メモリ32、演算部33、音声入出力装置34を備えている。また、4はダウンロード装置1から電話網を介してダウンロードされる音声認識ソフトウェアモジュールである。

【0017】図1を参照すると、サービス利用者から電話網2を通して通話があると回線制御部(NCU)15を介して制御部14に接続される。制御部14は音声認識ホスト部11に接続し、音声認識ホスト部11またはデータベース13から音声認識ソフトウェアモジュール4を、サービス利用者の汎用演算器つき電話端末3のメモリ32にダウンロードする。

【0018】電話端末3にダウンロードされた音声認識ソフトウェアモジュールは、演算部33により実行されると、まず電話端末3の端末種別、サービス利用者のいる環境雑音レベルなどを調べ、またこのサービスが話者を特定できるようなサービスの場合、サービス利用者の話者性(発話タイミング、発生音量、訛り等)なども、ダウンロードされた音声認識ソフトウェアモジュール4のデータ部42からセットし、サービス利用者の環境に最適になるように設定する。

【0019】音声認識するための準備が整うと、話者に発話を促すように信号音または音声再生され、話者は、これに従い音声入力装置34を介してサービスを利用する。電話端末3で音声認識された認識結果は局側の音声認識ホスト部11で規定された形式(API;アプリケーションインタフェース)に基づき転送され、音声認識ホスト部11は該認識結果に対応したサービスプログラムを実行し、音声応答部12は音声認識ホスト部11からの要求に応じて音声応答にて加入者にサービスを供給する。

【0020】本実施例においては、電話端末側の音声認識ソフトウェアモジュールが、局側から供給されたため、常に局側の認識ホスト部11のプログラムバージョンや、仕様との整合性をとることができると共に、サービス変更は、局側のダウンロード装置に内蔵されている音声認識ソフトウェアモジュールと音声認識ホスト部11のプログラムを入れ替えることで容易に行うことができる。

【0021】次に、本発明の第2の実施例について説明する。図2は、本発明の第2の実施例の構成を示すブロック図である。図2を参照すると、本発明の第2の実施例において、ダウンロード用音声認識ソフトウェアモジュール4のデータ部42には、予め既知の音響データ等

(回線の劣化を調べられるようなテストパターン)を格納しておき、ダウンロード用音声認識モジュール4が電話端末3で受信された時点で、このテストパターンを用いて回線の劣化状態を調べ、回線特性を判定することで、認識率向上に役立てるようにしたものである。

【0022】また、音声認識ホスト部11と既知の音響データを電話網2を通して通信し、音声認識率向上に役立てるようにしてもよい。この場合、ダウンロードされる音声認識ソフトウェアモジュール4は、通信モジュールとしての働きも有する。

【0023】また本発明の実施例においては、汎用演算器付き電話端末3にダウンロードされるソフトウェアモジュールとしては、音声認識用ソフトウェアに限らず、その他のサービスを実行するためのモジュールをダウンロードして電話端末側で実行させるようにしてもよい。

【0024】

【発明の効果】以上説明したように、本発明によれば、音声認識ソフトウェアモジュールを電話端末にダウンロードする構成としたことにより、端末型の音声認識性能を備えたと共に、さらに局側での柔軟なサービス変更が出来るシステムを構築することができる、という効果を奏する。

【0025】その理由は、局側が提供するいわゆる局用電話回線音声認識装置において、音声認識部を端末側に置くことにより、回線の雑音等による音声認識率の低下を防ぐと共に、端末側の音声認識ソフトウェアモジュールが局側から供給されるため、局側の認識ホスト部のプログラムバージョンや、仕様との整合性をとることができ、サービス変更は、局側のダウンロード装置に内蔵されている音声認識ソフトウェアモジュールと音声認識ホスト部のプログラムを入れ替えることによって行うことができるためである。

【図面の簡単な説明】

【図1】本発明の第1の実施例の構成を示す図である。

【図2】本発明の第1の実施例の構成を示す図である。

【図3】従来の局(センター)型の電話回線用音声認識装置の構成の一例を示す図である。

【図4】従来の端末型の電話回線音声認識装置の構成の一例を示す図である。

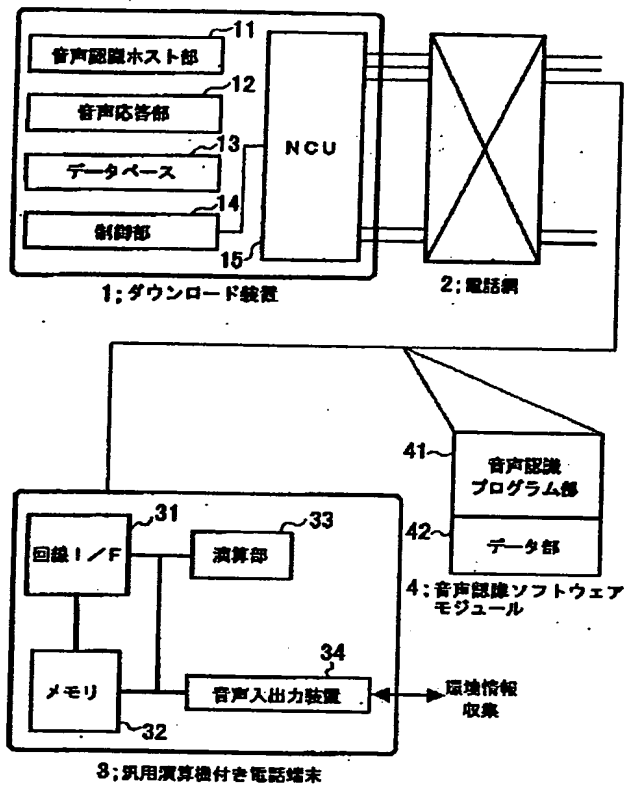
【符号の説明】

- 1 ダウンロード装置
- 11 音声認識ホスト部
- 12 音声応答部
- 13 データベース
- 14 制御部
- 15 回線制御ユニット(NCU)
- 2 電話網
- 3 汎用演算機付き電話端末
- 31 回線I/F(インタフェース)
- 32 メモリ

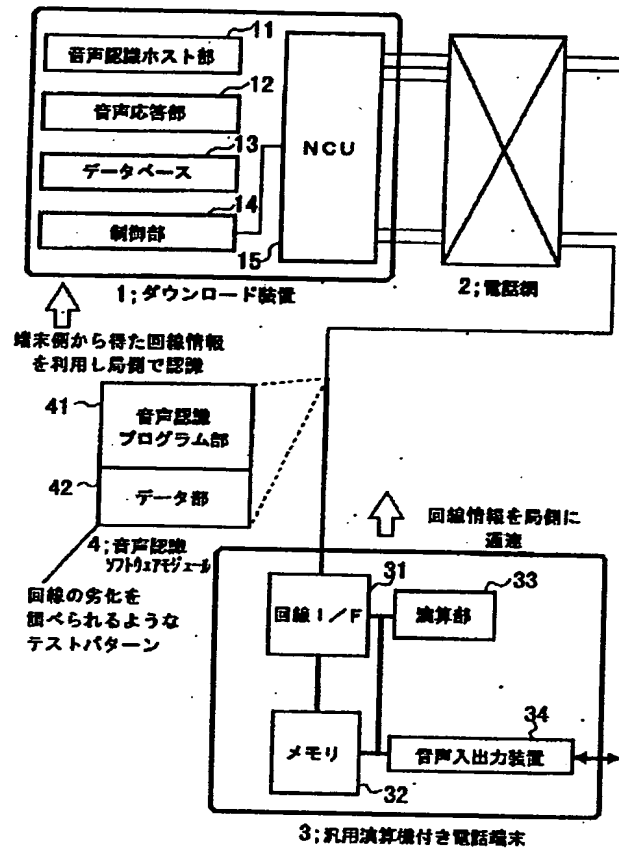
- 33 演算部
- 34 音声入出力装置
- 4 音声認識ソフトウェアモジュール

- 41 音声認識プログラム部
- 42 データ部

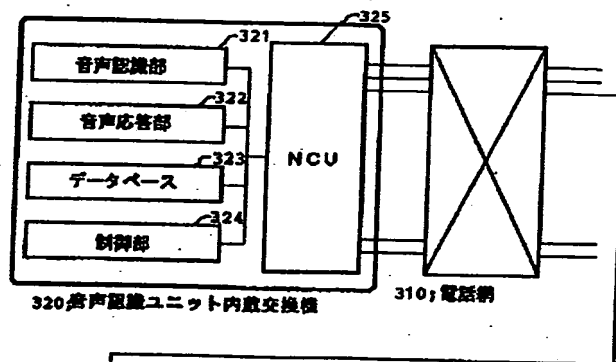
【図1】



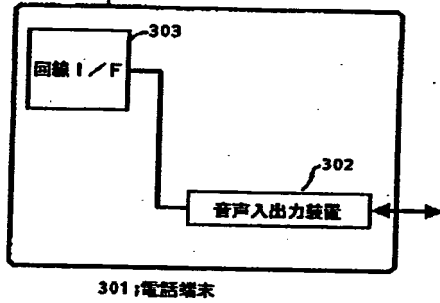
【図2】



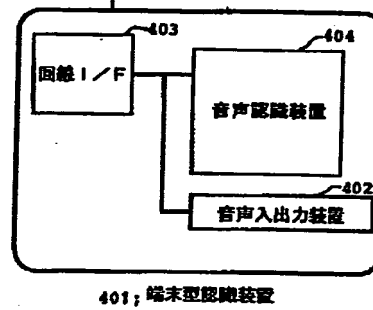
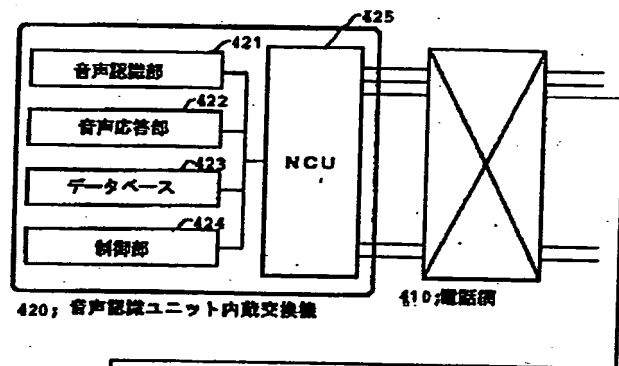
【図3】



⇔  
回線で音声が悪化



【図4】



認識機能は始め  
からは端末側に  
固定されている  
ため、局側と連携  
したサービス変更  
が困難